Euphytica (2008) 162:291–300
DOI 10.1007/s10681-008-9675-8

Discovery and detection of single nucleotide polymorphism (SNP) in coding and genomic sequences in chickpea (*Cicer arietinum* L.)

P. N. Rajesh · Fred J. Muehlbauer

Received: 26 October 2007 / Accepted: 14 February 2008 / Published online: 26 February 2008

Abstract Chickpea genetic mapping has been hampered by insufficient amplicon length polymorphism for sequence based markers. To develop an alternative source of polymorphic markers, we determined naturally abundant single nucleotide polymorphism (SNP) in coding and genomic regions between FLIP 84-92C (*C. arietinum*) and PI 599072 (*C. reticulatum*) and identified an inexpensive method to detect SNP for mapping. In coding sequences, 110 single base changes or substitutions (47% transitions and 53% transversions) and 18 indels were found; while 50 single base changes (68% transitions and 33% transversions) and eight indels were observed in genomic sequences. SNP frequency in coding and genomic regions was 1 in 66 bp and 1 in 71 bp, respectively. In order to effectively use this high frequency of polymorphism, we used Cleaved Amplified Polymorphic Site (CAPS) and derived CAPS (dCAPS) marker systems to identify a restriction site at SNP loci. In this study, we developed six CAPS and dCAPS markers and fine mapped *QTL*1, a region previously identified as important for ascochyta blight resistance. One of the CAPS markers from a BAC end was identified to account for 56% of the variation for ascochyta blight resistance in chickpea. Conversion of naturally abundant SNPs to CAPS and dCAPS for chickpea mapping, where absence of amplicon length polymorphism is a constraint, has potential to generate high-density maps necessary for map-based cloning and integration of physical and genetic maps.

Keywords Ascochyta blight · Indels · Single nucleotide polymorphism · Transition · Transversion

P. N. Rajesh (✉)
Department of Crop and Soil Sciences, Washington State University, 303 Johnson Hall, Pullman, WA 99164-6434, USA
e-mail: pnrajesh@wsu.edu

F. J. Muehlbauer
Department of Crop and Soil Sciences, USDA-ARS, Washington State University, 303 Johnson Hall, Pullman, WA 99164-6434, USA
e-mail: muehlbau@wsu.edu

Introduction

Chickpea is the third most important pulse crop worldwide. World production is 8.6 million metric tons annually from 11.2 million hectares (FAOSTAT data 2005). The genome size of chickpea is estimated to be 740 Mb (Arumuganathan and Earle 1991) which is 1.5 times that of *Medicago truncatula*, a model legume. There are six published linkage maps that were generated using morphological and various molecular markers such as isozymes, Amplified Fragment Length Polymorphism (AFLP), Random Amplified Polymorphic DNA (RAPD), Inter Simple Sequence Repeat (ISSR), Resistant Gene Analog (RGA) and Sequence Tagged Microsatellite Site

(STMS) markers (Winter et al. 2000; Santra et al. 2000; Cho et al. 2002, 2004; Flandez-Galvez et al. 2003; Collard et al. 2003). In these linkage maps, many agronomically important traits do not have tightly linked markers for use in marker assisted selection or map based cloning. There were large gaps in all linkage maps due to an insufficient number of markers needed to fill the gaps. In order to generate a linkage map without large gaps, additional markers need to be developed and mapped. STMS markers (Hüttel et al. 1999; Winter et al. 1999) were 78% polymorphic in the FLIP 84-92C × PI599072 cross and were used to increase marker density of the genetic map that improved our understanding of the chickpea genome (Tekeoglu et al. 2002). However, the number of STMS markers currently available is limited at approximately 600 (Hüttel et al. 1999; Lichtenzveig et al. 2005; Sethy et al. 2006). In addition, the availability of sequence based markers was also limited by the number of sequences deposited in the public databases.

Chickpea genomic research has progressed rapidly over the past 3–4 years as evidenced by the availability of genomic tools such as BAC libraries (Rajesh et al. 2004; Lichtenzveig et al. 2005), cDNA libraries (Buhariwalla et al. 2005; Coram and Pang 2005), availability of BAC sequences (734,456 bp) (http://www.genome.ou.edu/plants_totals.html), 48,339 BAC end sequences and 1,960 Expressed sequence tags (ESTs) in the public database http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=nucleotide&cmd=search&term=chickpea) as of 23 October 2007. Since there are many gene and locus specific sequence based markers available currently, there is an urgent need for an alternative source for polymorphism in order to map these markers that will increase the marker density at important genomic regions and also will facilitate cloning of genes of interest in chickpea.

Single DNA base differences such as transitions (change of a purine for a purine or a pyrimidine for a pyrimidine) and transversions (change of a purine for pyrimidine or a pyrimidine for a pyrimidine) between homologous DNA fragments plus small insertions and deletions (indels), collectively referred to as single-nucleotide polymorphisms (SNPs) (Zhu et al. 2003). They are di, tri and tetra allelic in nature although diallelic is common in humans (Kahl et al. 2004). SNPs have a wide variety of applications such as association mapping, marker density increase, QTL mapping, high throughput marker assisted selection and evolutionary biology (Lopez et al. 2005). Depending on their location in the genome and impact on expression of a gene, they are classified into non-coding SNPs (ncSNPs), coding SNPs, exonic SNPs, cDNA SNPs and candidate SNPs (Kahl et al. 2004). SNPs in non-coding and coding regions have been extensively studied only in few plant species such as *Arabidopsis* (Jander et al. 2002), maize (Ching et al. 2002), and soybean (Zhu et al. 2003). There have been no reports of SNP discovery and detection to understand the chickpea genome.

Ascochyta blight (caused by *Ascochyta rabiei* (Pass). Lab.) is a devastating disease worldwide and has the potential to cause up to 100% yield loss to chickpea. Although earlier genetic studies identified two *QTL* (*QTL*1 and *QTL*2) that confer resistance to this disease, the markers flanking these *QTL*s were dominant RAPD markers and were quite distant from each other. For example, *QTL*1 in linkage group VIII which reportedly accounted for 35% of the variation in ascochyta blight resistance was flanked by UBC733b and UBC181a at a genetic distance of 11 cM (Tekeoglu et al. 2002). In order to determine the number of genes contributing to ascochyta blight resistance at *QTL*1 and narrow the genetic distance between the flanking markers, it is necessary to increase the marker density at this genomic region. We constructed a *Hind*III BAC library from FLIP84-92C cultivar to generate new sequence based markers and also for gene discovery (Rajesh et al. 2004).

In this study, our objectives were: i. SNP discovery in coding sequences and genomic regions; ii. Determine a cost-effective SNP detection method; and iii. To utilize SNPs for genetic mapping. Since ESTs for chickpea are available in the database, SNP discovery will assist in gene mapping and marker density increase in linkage maps. To discover SNPs in genomic sequences, we chose two genomic regions that are agronomically important because genes for ascochyta blight resistance and fusarium wilt resistance were mapped to those regions (Tekeoglu et al. 2002; Sharma et al. 2004). There are many expensive methods available for SNP discovery such as pyrosequencing (Alderborn et al. 2000; Ching and Rafalski 2002), mass spectrometry (Rodi et al. 2002) and Affymetrix chips (Borevitz et al. 2003) and for SNP detection such as primer extension, allele specific

oligonucleotide hybridization and allele specific oligonucleotide ligation (Kahl et al. 2004). However, a combination of allele sequencing for SNP discovery and generation of cleaved amplified polymorphic sequence (CAPS) and derived CAPS (dCAPS) markers at SNP loci for SNP detection is a relatively inexpensive method. CAPS markers have been used efficiently in Arabidopsis (Konieczny and Ausubel 1993), melon (Morales et al. 2004) and pea (Konovalov et al. 2005). Generation of genome wide CAPS markers has been used successfully in developing a linkage map for *M. truncatula*, a model legume (Choi et al. 2004). Here, we chose *QTL*1 for ascochyta blight resistance (Tekeoglu et al. 2002) as a target genomic region to increase marker density by converting SNP loci at BAC ends and a gene identified from this region into CAPS and dCAPS markers.

Materials and methods

Plant materials

Chickpea accession FLIP84-92C and wild relative PI599072 that are resistant and susceptible to *Ascochyta rabiei* (Pass). Lab., respectively, were used in this study. The F_2 population from this cross was advanced by single seed descent to the F_8 to produce F_7 derived RILs (Santra et al. 2000; Tekeoglu et al. 2002). A population of 142 $F_{7:8}$ derived recombinant inbred lines (RILs) from the FLIP 84-92C (resistant) × *C. reticulatum* PI 599072 (susceptible) cross was used for genetic mapping. DNA was extracted from leaf tissue of each RIL and the parental lines according to Doyle and Doyle (1987).

PCR conditions for amplification of ESTs and genomic sequences

The primers for ESTs and BAC ends were designed using http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi and synthesized by MWG Biotech, USA. PCR amplification using these primers was performed in 10 mM Tris–HCl pH 8.3, 50 mM KCl, 0.1% TritonX100, 2.5 mM $MgCl_2$, 0.2 μM dNTP, 50 pmole of forward and reverse primers, 1 unit of Taq polymerase (Promega, USA) and approximately 20–40 ng of genomic DNA per 20 μl reaction. The amplification was carried out using a Gene Amp PCR system 9700 (Perkin Elmer, USA). After initial denaturation at 94°C for 2 min, 40 cycles of PCR were carried out. Each cycle was comprised of 20 s denaturation at 94°C, 50 s annealing at a different temperature depending on the primers (Tables 1, 2) and 2 min extension at 72°C for 50 s followed by final extension at 72°C for 8 min. The amplified products were separated on 2% agarose gel and stained with ethidium bromide.

PCR product sequencing

EST primers (Table 1) used in this study were obtained from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=nucleotide.

Direct PCR sequencing was performed for all ESTs. In this method, genomic DNA was amplified using respective EST primers as described in "PCR conditions for amplification of ESTs and genomic sequences" section and 2 μl of the PCR product was run on a 2% agarose gel to ensure that the product was a single band. Each amplified product was then treated with EXOSAP-IT enzyme (USB, Cleveland, OH) following manufacturer's instruction. Sequencing of the treated samples was performed on an ABI Prism 377 DNA sequencer (Applied Biosystems, USA) using the dideoxy sequencing method.

BAC end sequencing

BAC DNAs from FLIP84-92C BAC library (Rajesh et al. 2004) were purified using the BAC_{96} Miniprep kit (Millipore, USA). The primers for end sequencing were designed from the flanking region of the *Hind*III cloning site and synthesized by MWG Biotech, USA. The sequences of the primers are: V41F-GCGATT AAGTTGGGTAACGC; V41R-CACAGGAAACAG CTATGACC. One and a half micrograms of clean BAC DNA was used for direct PCR cycle sequencing with extended (66) cycles on an ABI PRISM 377 automatic sequencer (Applied Biosystems, USA).

SNP discovery and CAPS and dCAPS marker development

Markers were developed from the ends of the BACs that had been identified by super-pool method described in Rajesh et al. (2004) and hybridization of chickpea BAC library using OPS06-1 and SCAR733b, respectively. SNPs were discovered by

Table 1 Description of ESTs and summary of detected SNPs between *C. arietinum* (FLIP84-92C) and *C. reticulatum* (PI 599072)

NCBI name	Gene name	Oligos	Ann. temp	Sequence (5′–3′)	Amplicon size (bp)	Indel	Transistion	Transversion
AJ746343	Esterase (est gene) F	est F	55	tcccatcaatgttctcacca	600	0	0	0
	Esterase (est gene) R	est R		gttcaaatgatgatggcaaaga				
AY635930	Kunitz proteinase inhibitor-1 (kpi1 gene) F	kpi1 F	55	tccaggtggaggagtcagat	351	3	3	2
	Kunitz proteinase inhibitor-1 (kpi1 gene) R	kpi1 R		tcaaacgtcttccaccttca				
AJ276262	Trypsin protein inhibitor F	tpi2 F	59	tccagcaagtgataaccctaaaac	317	2	2	3
	Trypsin protein inhibitor R	tpi2 R		gacgcatcaaatccaccttt				
AJ609275	Leucine rich repeat protein (lrr gene) F	lrr F	55	ttccttccaccattggaaac	662	0	2	2
	Leucine rich repeat protein (lrr gene) R	lrr R		taattcactttcgggtcacg				
AJ489612	Nitrate transporter (ntp gene) F	ntp F	55	caaaggattgccattggact	425	2	1	2
	Nitrate transporter (ntp gene) R	ntp R		cccatttgcattgatggctta				
AJ012690	Chalcone synthase F	chs F	59	gcaagacatggtggttgttg	282	1	0	4
	Chalcone synthase R	chs R		cacttccctagttgccctca				
AJ131047	Glucan-endo-1,3-beta-glucosidase F	ge13bg F	55	tccaattggtggctcttctt	590	4	33	30
	Glucan-endo-1,3-beta-glucosidase R	ge13bg R		ccaccaaatccaaacggata				
AB025004	Transketolase F	tk F	59	ggacgctacaaggcagaatc	660 (471)[a]	6	1-intron; 8-exon	1-intron; 10-exon
	Transketolase R	tk R		acattgcgctcttctggagt				
AF416481	Serine hydroxy methyl transferase-F	SHMT-F	55	gaagatgtctgatcaagttttgc	554 (284)[a]	0	1	2-intron; 1-exon
	Serine hydroxy methyl transferase-R	SHMT-R		tgcaaaacaatttccaacca				
AF416480	Aldolase-F	Aldol-F	59	cccaagaagggaattgtttg	180	0	1	1
	Aldolase-R	Aldol-R		aagttgtaaagctgcatgtaaacc				
				Total: 4,621 bp	Average: 462	18	52	58

Ann. Temp, annealing temperature; amplicon size, aligned sequence length; indel, total number of indels; transistion, total number of all interchanges between purines and between pyramidines; transversion, total number of all interchanges between purines and pyramidines

[a] The values in parentheses are the amplified exon size

aligning sequences from parental DNAs using Vector NTI Advance 9.0 software (www.informaxinc.com). The output files of sequences were analyzed by ABI sequencing analysis 5.2 program and quality values were assigned by KB base calling procedure. The base quality value (QV) was visualized using ABI sequence scanner v1.0. Each SNP was confirmed by checking chromatogram visually also. We developed six CAPS and dCAPS (Table 2) from BAC ends and a gene and mapped the markers to the *QTL*1 region for ascochyta blight resistance. In CAPS analysis, the sequences of both parental DNAs were restriction mapped at SNP loci to identify a suitable restriction enzyme (Fig. 1). The primers for dCAPS analysis were designed using web based software package http://helix.wustl.edu/dcaps/dcaps.html (Fig. 2) with a single nucleotide mismatch adjacent to SNP position creating restriction site in the amplified PCR product of one parent but not the other (Neff et al. 1998). *HpyCH4*III, *Taq*1, *Afl*III, *Aci*I are the various restriction enzymes we used for mapping these markers.

Linkage analysis

Segregation of marker loci was tested for goodness of fit to the expected Mendelian ratio of 1:1 using χ^2 analysis ($P < 0.05$). Markers with distorted distribution were also used for linkage analysis. Linkage analysis was performed using Mapmaker/Exp 3.0 (Lander et al. 1987). Linkage groups were established at a constant LOD score of 4.0 and a recombination value of 0.25 by two point analyses using the 'group' command. These new markers were run along with the other previously scored marker data (Tekeoglu et al. 2002). The most likely order of loci within a group was determined using multipoint 'compare' command and these orders were verified using the 'ripple' command. The Kosambi mapping function was used to determine cM distances between markers (Kosambi 1944) (Fig. 3). Double crossovers were checked by 'double crossover' command in Map Manager QTb (version 2.8) (Manly 1998). QTL analysis for blight resistance was carried out with the simple interval mapping function using Qgene (Nelson 1997) at an LOD score of 3.0. Single-point regression analysis was used to identify markers significantly associated with blight resistance.

Results

SNP discovery

A total of 8,726 bp (4,621 from coding regions and 4,105 from genomic sequences) from each parental line, FLIP-84-92C and PI 599072, were analyzed for SNP discovery. The number of base pairs was calculated by adding the amplicon size that we obtained after sequencing as described in Tables 1, 2. To obtain sequences from the coding regions, we designed primers for 10 representative ESTs (Table 1) that are reported to play a role in defense response. These primers amplified 4,621 bp altogether that were used for SNP discovery in this study (Table 1).

Aldolase and serine hydroxyl methyl transferase (SHMT) were identified as ascochyta blight responsive genes in our earlier study (Rajesh et al. 2003). All of the primers designed from ESTs assessed amplified coding regions only except transketolase and SHMT. A total of 110 single base changes or substitutions (47% transitions and 53% transversions) and 18 indels were observed in expressed sequences analyzed in this study. Interestingly, a very high number of transition and transversion (33 and 30, respectively) was discovered in glucan-endo-1, 3-beta-glucosidase product when compared to other ESTs used in this study. Among the ESTs, one indel for every 257 bp, one nucleotide substitution for 42 bp and over all one SNP in 36 bp was observed. Excluding glucan-endo-1, 3-beta-glucosidase, overall one SNP in 66 bp was estimated to be present in the sequences under study.

In asochyta blight and fusarium wilt disease resistance genomic regions together, 50 single base changes (68% transitions and 33% transversions) and eight indels were observed. In the fusarium wilt resistance genomic region represented by AY356156 and AY356154 in Table 2, one indel in 106 bp, one nucleotide substitution in 38 bp and overall one SNP in 28 bp were estimated to be present. On the contrary, low level of SNP was detetcted at the ascochyta blight resistance genomic region represented by AC161102, AC145766, AC145456, AC161104 and AC161105 in Table 2 which corresponds to indels of 1 in 1,192 bp, nucleotide substitution of 1 in 99 bp and cumulatively we estimated that there was an average of one SNP in

Table 2 Description of genomic sequences and summary of detected SNPs between FLIP 84-92C and PI599072

NCBI name	Gene name	Oligos	Type[a]	Ann. temp	Sequence (5′–3′)	Amplicon size (bp)	Indel	Transition	Transversion
AY356156	Resistance gene region F	rgr2 F	Locus specific	59	cagccctttatgcctgatcc	370	3	1	6
	Resistance gene region R	rgr2 R			aagccttggtttcttggacag				
AY356154	Resistance gene region F	rgr4 F	Locus specific	59	actcaccaccatgtcccttc	160	2	3	4
	Resistance gene region R	rgr4 R			aaaattcagccatatgtcaaacc				
AC161102	BAC clone	LF-F	Scar	55	cggcgtgaaagtgttgttta	239	0	2	0
		LF-R			tgaatccatgcgagtcagaa				
AC161102	BAC clone	PNR2F	Scar	59	agggagagtagataggtcaac	688	0	1	0
		PNR2R			agggagcaaataataatacgtattt				
AC161102	BAC clone	LR-F	CAPS-*HpyCH4III*	55	tctcaccttcaagggtttcc	341	0	2	0
		LR-R			tttcagcaaaagcgacattg				
AC145766	BAC clone	MF-F	dCAPS-*Taq1*	59	aacttgaagatatttaatatggcactc	553	0	2	0
		MF-R			ataccccccacagaccaagg				
AC145766	BAC clone	SR-2L	CAPS-*HpyCH4III*	35	ggagagcatggagactcaaaaa	414	0	9	1
		SR-2R			cggctaaacctagtcgttcaaa				
AC145456	BAC clone	150F-F	dCAPS-*AciI*	55	gcctgcaggaatagatctacagc	202	0	3	1
		150F-R			agcacctaatggcaccactcc				
AC145456	BAC clone	UF3T-F	dCAPS-*AflIII*	55	tttgaaaggacttcttccacatg	237	0	3	1
		UF3T-R			aatggcctgcaaatcatagg				
AC145456	BAC clone	150R-F	CAPS-*AciI*	55	gcactgagctttcctgaacc	241	0	4	0
		150R-R			tgcgattagggttgagttgaac				
AC161104	BAC clone	CP-9IK1F-F	Locus specific	59	aagaggacaggctgaagcaa	355	2	2	1
		CP-9IK1F-R			ccttctttcccatctcctaca				
AC161105	BAC clone	CP-9IN1F-F	Locus specific	59	gcaaggcattttcattagcc	305	1	2	2
		CP-9IN1F-R			accccaagtccttaggtgct				
					Total: 4,105 bp	Average: 342	8	34	16

[a] Type: Locus specific markers (rgr2, rgr4, CP-9IK1, CP-9IN1) were not used for mapping; CAPS and dCAPS markers that were designed to create polymorphism at monomorphic loci using SNPs and scar-sequence characterized amplified region were used for genetic mapping in this study

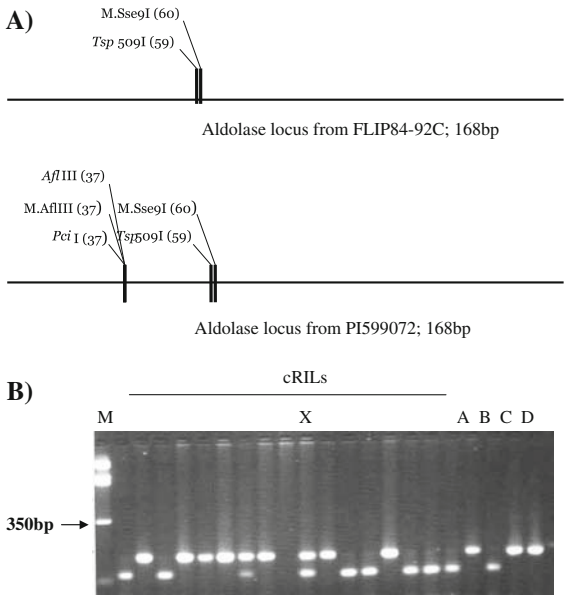

Fig. 1 CAPS analysis using Aldolase specific primers between FLIP 84-92C and PI599072. (**a**) Sequence schematic; A naturally present SNP at 37th base pair in PI599072 created *Afl*III restriction site which created polymorphism between parental lines. (**b**) The amplified products were digested with *Afl*III and ran on 2% agarose gel. A—*Afl*III digested FLIP 84-92C, B—*Afl*III digested PI599072, C—Undigested FLIP84-92C, X—Codominant allele, D—Undigested PI599072, M—Lambda Bst-N1 marker, cRILs—Chickpea recombinant inbred lines

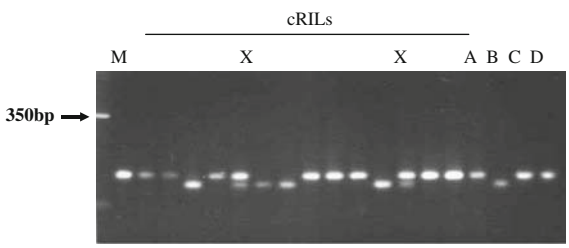

Fig. 2 dCAPS analysis (UF3T) using BAC end primers between FLIP 84-92C and PI599072. dCAPS primer for UF3T (TTTGAAAGGACTTCTTCCA**C**ATG) was designed by incorporating a nucleotide (C) adjacent to SNP that created *Afl*III enzyme site (ACRYGT) in PI599072 (TTTGAAAGGACTTCTTCCA**C**ATG**T**GTTT) but not in FLIP84-92C (TTTGAAAGGACTTCTTCCAAATGGGTTT). The amplified products were then digested with *Afl*III enzyme and ran on 2% agarose gel. A—*Afl*III digested FLIP 84-92C, B—*Afl*III digested PI599072, C—Undigested FLIP84-92C, X—Codominant allele, D—Undigested PI599072, M—Lambda Bst-N1 marker, cRILs—chickpea recombinant inbred lines

every 92 bp. Including both disease resistance genomic regions, it was calculated that these regions harbor the following: indels—1 in 513 bp; nucleotide

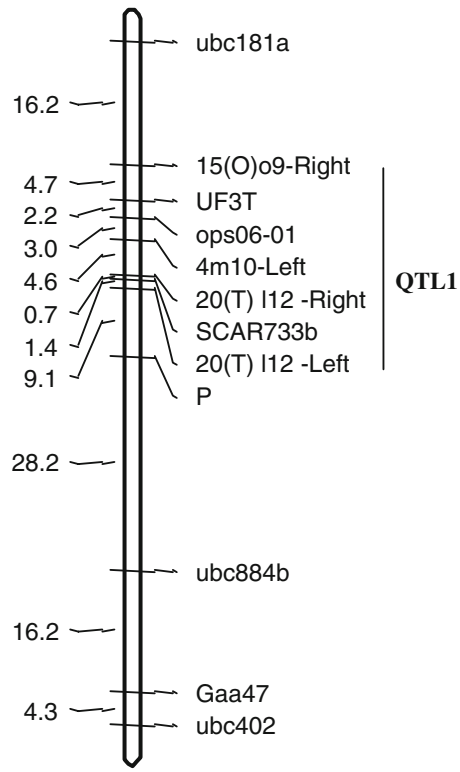

Fig. 3 The linkage group with an increased marker density at *QTL*1 for Ascochyta blight resistance developed using 142 $F_{7:8}$ RILs from the interspecific cross of the cultivated chickpea line FLIP 84-92C (resistant) × *C. reticulatum* PI599072 (susceptible). Our linkage group is same as the linkage group VIII of Tekeoglu et al. (2002). Genetic distance in cM are to the left of the linkage group

substitution—1 in 82 bp and over all one SNP per 71 bp. Transition was estimated to be higher in genomic region (68%) than the coding region (47%) while transversion (53%) was more in coding region when compared to genomic region (33%).

We calculated the nucleotide diversity (Π) using the following formula: Total number of SNPs/total length of aligned sequences. Based on this formula, we estimated the nucleotide diversity for ESTs as 0.028 and for genomic regions as 0.014.

Mapping SNPs

To enrich the marker density at *QTL*1 region, as a first step, we converted UBC733b into a single locus Sequence Characterized Amplified Region marker (SCAR733b). We screened our BAC library with OPS06-01 (Rakshit et al. 2003) and SCAR733b, two

markers that are in the flanking regions of *QTL*1. We generated markers from the ends of identified BAC clones and a gene present in one of the BACs. Since these PCR based markers are monomorphic in amplicon length between the parental lines (FLIP84-92C and PI599072), we identified Cleaved amplified polymorphic site (CAPS) and derived CAPS (dCAPS) marker systems as a suitable method by utilizing the SNPs to increase marker density at *QTL*1 (Fig. 3).

For CAPS, the sequences from different parental lines were aligned and restriction sites present at SNP loci were identified. SNPs that conferred differential restriction enzyme sites between the parents were used for further analysis (Fig. 1).

At some SNP loci, a suitable enzyme was not identified that would cleave the amplification product. In such cases, the dCAPS technique was used by designing primers with a single nucleotide mismatch adjacent to the SNP position and creating restriction site in the amplified PCR product of one parent but not the other (Fig. 2).

Discussion

SNP analysis

Introns, 5′ and 3′ untranslated regions and BAC ends are suggested to possess more SNPs than exons (Rafalski 2002) because this change will have minimal effect on the expressed product. Although all our EST products amplified only the coding regions, transketolase and SHMT amplified introns as well which was evidenced by amplification of larger amplicon sizes than expected. In Table 2, the sizes of the exon in parentheses which was expected and the total amplicon of transketolase and SHMT are mentioned. Interestingly, a higher frequency of SNPs was found in exons than in introns of transketolase. No indels were detected in introns of transketolase and SHMT.

Rgr2 and Rgr4 are ends of BACs identified using a STMS marker Ta96 which has been reported to be linked to the *Fusarium oxysporum* f. sp. *ciceris* race 3 (*Foc*3) resistance gene (Sharma et al. 2004). Winter et al. (2000) mapped many STMS markers on linkage group II where *Foc*3, *Foc*4 and *Foc*5 were located which indicated that microsatellites are abundant in this genomic region. A high SNP frequency was estimated at *Foc*3 resistance genomic region and it could be due to "mutation active" nature of this part of the genome.

Our estimated SNP frequency of 1 in 66 bp in coding sequences excluding glucan-endo-1, 3-beta-glucosidase and 1 in 71 bp in genomic regions in chickpea is higher than estimates for wheat [1 in 1,000 bp (Bryan et al. 1999)], human [1 in 1,000 bp] (Sachidanandam et al. 2001) and *Arabidopsis* [1 in 3.3 kb] (Jander et al. 2002). However, SNP frequency in chickpea is comparable to sugar beet which has 1 in every 126 bp (Schneider et al. 2001), rice with 1 in 89 bp (average among five rice genotypes), (Nasu et al. 2002), maize that has 1 SNP per 60–120 bp (Ching et al. 2002), cassava with 1 in 62 bp (Lopez et al. 2005), and pearl millet (one in every 59 bp in introns and less in exons) (Bertin et al. 2005).

Nucleotide diversity (Π) in chickpea (0.028 in coding and 0.014 in genomic regions) is similar to maize which was estimated to be 0.016 (Ching et al. 2002). High SNP density in chickpea genome identifies the alternative resource for determining polymorphism.

Mapping

We chose FLIP-84-92C and PI 599072 for SNP discovery and genetic mapping because recombinant inbred lines (RILs) developed from crossing these two parental lines were previously used to generate linkage maps (Santra et al. 2000; Tekeoglu et al. 2002; Rakshit et al. 2003). For genetic mapping, we adopted sequencing of an allele from genotypes to discover SNPs, identified visually the sequence variation and detected them by designing CAPS or dCAPS markers. Most of these markers are co-dominant in nature and hence we can differentiate homozygous and heterozygous conditions. This method is inexpensive and requires simple agarose gel electrophoresis, ethidium bromide staining and visualization under UV similar to other quality gels.

Mapping of SNPs by developing CAPS and dCAPS markers increased marker density at *QTL*1 and narrowed the genetic distance from 11 cM to 0.7 cM. Marker-trait association using QGene software program discovered one of the BAC ends which had been converted into a CAPS marker

Table 3 Statistics of newly identified markers flanking *QTL*1 for ascochyta blight resistance

Marker	R^2 in %	LOD	Genetic distance
20(T)l12-Right	56	19.98	0.7 cM
SCAR733b	51	18.04	1.4 cM
20(T)l12-Left	48	15.62	

R^2 and LOD value of these markers were calculated using QGene software

$P < 0.0001$

(20(T)l12-Right) that accounted 56% of the variation in ascochyta blight resistance at LOD value 19.98, an improvement of the previously reported (35% and LOD 13.40) (Table 3).

In conclusion, we discovered SNPs in coding and genomic regions and detected those natural nucleotide polymorphism to increase marker density within an agronomically important genomic region. In genome sequencing, the new strategies are focusing more on sequencing gene rich regions rather than sequencing the entire genome which has more than 70% repetitive sequences. Similarly, in constructing linkage maps, mapping ESTs, BAC ends and other locus specific markers spanning genomic regions where important traits are present are given higher priority than random markers. Our study discovered a high frequency of SNPs in the chickpea genome that can be used for mapping sequence based markers. This discovery promises to impact mapping efforts by the chickpea research community.

Acknowledgements PNR wishes to thank McKnight Foundation, USA for financial support. The authors thank Jackie Corona and Sheri Babb for their help in editing this manuscript.

References

Alderborn A, Kristofferson A, Hammerling U (2000) Determination of single-nucloetide polymorphisms by real-time pyrophosphate DNA sequencing. Genome Res 10:1249–1258

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9:208–218

Bertin I, Zhu JH, Gale MD (2005) SSCP-SNP in pearl millet—a new marker system for comparative genetics. Theor Appl Genet 110:1467–1472

Borevitz JO, Liang D, Plouffe D, Chang HS, Zhu T, Weigel D, Berry CC, Winzeler E, Chory J (2003) Large-scale identification of single-feature polymorphisms in complex genomes. Genome Res 13:513–523

Bryan GJ, Stephenson P, Collins A, Kirby J, Smith JB, Gale MD (1999) Low levels of DNA sequence variation among adapted genotypes of hexaploid wheat. Theor Appl Genet 99:192–198

Buhariwalla HK, Jayashree B, Eshwar K, Crouch JH (2005) Development of ESTs from chickpea roots and their use in diversity analysis of the *Cicer* genus. BMC Plant Biol 17:5–16

Ching A, Rafalski A (2002) Rapid genetic mapping of ESTs using SNP pyrosequencing and indel analysis. Cell Mol Biol Lett 7:803–810

Ching ADA, Caldwell KS, Jung M, Dolan M, Smith OS, Tingey S, Morgante M, Rafalski A (2002) SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. BMC Genet 3:19

Cho S, Kumar J, Shultz JL, Anupama K, Tefera F, Muehlbauer FJ (2002) Mapping genes for double podding and other morphological traits in chickpea. Euphytica 128:285–292

Cho S, Chen W, Muehlbauer FJ (2004) Pathotype-specific genetic factors in chickpea (*Cicer arietinum* L.) for quantitative resistance to ascochyta blight. Theor Appl Genet 109:733–739

Choi HK, Kim D, Uhm T, Limpens E, Lim H, Mun JH, Kalo P, Penmetsa RV, Seres A, Kulikova O, Roe B, Bisseling T, Kiss GB, Cook DR (2004) A sequence-based genetic map of *Medicago truncatula* and comparison of marker colinearity with *M. sativa*. Genetics 166:1463–1502

Collard BC, Pang EC, Ades PK, Taylor PW (2003) Preliminary investigation of QTLs associated with seedling resistance to ascochyta blight from *Cicer echinospermum*, a wild relative of chickpea. Theor Appl Genet 107:719–729

Corum T, Pang ECK (2005) Isolation and analysis of candidate ascochyta blight defence genes in chickpea. Part I. Generation and analysis of an expressed sequence tag (EST) library. Physiol Mol Plant Pathol 66:192–200

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small amount of fresh leaf tissue. Phytochem Bull 19:11–15

FAOSTAT (2005) http://faostat.fao.org/faostat/collections?subset=agriculture

Flandez-Galvez H, Ford R, Pang ECK, Taylor PWJ (2003) An intraspecific linkage map of the chickpea (*Cicer arietinum* L.) genome based on sequence tagged microsatellite site and resistance gene analog markers. Theor Appl Genet 106:1447–1456

Hüttel B, Winter P, Weising K, Choumane W, Weigand F, Kahl G (1999) Sequence-tagged microsatellite-site markers for chickpea (*Cicer arietinum* L.). Genome 42:210–217

Jander G, Norris SR, Rounsley SD. Bush DF, Levin IM, Last RL (2002) *Arabidopsis* map-based cloning in the post-genome era. Plant Physiol 129:440–450

Kahl G, Mast A, Toke N, Shen R, Boom DVD (2004) SNPs: detection techniques and their potential for genotyping and genome mapping. In: The handbook of plant genome mapping. Wiley-VCH Verlag GmbH & Co, pp 75–107

Konieczny A, Ausubel FM (1993) A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. Plant J 4:403–410

Konovalov F, Toshchakova E, Gostimsky S (2005) A caps marker set for mapping in linkage group iii of pea (*Pisum sativum* l.) Cell Mol Biol Lett 10:163–171

Kosambi DD (1944) The estimation of map distance from recombination values. Ann Eugen 12:172–175

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Lichtenzveig J, Scheuring C, Dodge J, Abbo S, Zhang HB (2005) Construction of BAC and BIBAC libraries and their applications for generation of SSR markers for genome analysis of chickpea, *Cicer arietinum* L. Theor Appl Genet 110:492–510

Lopez C, Pie'gu B, Cooke R, Delseny M, Tohme J, Verdier V (2005) Using cDNA and genomic sequences as tools to develop SNP strategies in cassava (*Manihot esculenta* Crantz). Theor Appl Genet 110:425–431

Manly KF (1998) User's manual for map manager classic and map manager QT

Morales M, Roig E, Monforte AJ, Arús P, Garcia-Mas J (2004) Single-nucleotide polymorphisms detected in expressed sequence tags of melon (*Cucumis melo* L.). Genome 47:352–360

Nasu S, Suzuki J, Ohta R, Hasegawa K, Yui R, Kitazawa N, Monna L, Minobe Y (2002) Search for and analysis of single nucleotide polymorphism (SNPs) in rice (*Oryza sativa*, O*ryza rufipogon*) and establishment of SNP markers. DNA Res 9:163–171

Neff M, Neff J, Chory J, Pepper A (1998) dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in *Arabidopsis thalaiana* genetics. Plant J 14:387–392

Nelson JC (1997) QGENE: software for maker-based genomic analysis and breeding. Mol Breed 3:239–245

Rafalski A (2002) Applications of single nucleotide polymorphisms in crop genetics. Curr Opin Plant Biol 5:94–100

Rajesh PN, Gupta VS, Ranjekar PK, Muehlbauer FJ (2003) Functional genome analysis using DDRT with respect to ascochyta blight disease in chickpea. Int Chickpea Pigeonpea Newsl 10:35–37

Rajesh PN, Coyne C, Meksem K, Sharma KD, Gupta V, Muehlbauer FJ (2004) Construction of a *Hind*III bacterial artificial chromosome library and its use in identification of clones associated with disease resistance in chickpea. Theor Appl Genet 108:663–669

Rakshit S, Winter P, Tekeoglu M, Juarez Muñoz J, Pfaff T, Benko-Iseppon AM, Muehlbauer FJ, Kahl G (2003) DAF marker tightly linked to a major locus for *Ascochyta* blight resistance in chickpea (*Cicer arietinum L.*). Euphytica 132:23–30

Rodi CP, Darnhofer-Patel B, Stanssens P, Zabeau M, van den Boom D (2002) A strategy for the rapid discovery of disease markers using the MassARRAY system. BioTechniques 32:62–69

Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, Willey DL et al (2001) A map of human genome. sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409:928–933

Santra DK, Tekeoglu M, Ratnaparkhe MB, Gupta VS, Ranjekar PK, Muehlbauer FJ (2000) Identification and mapping of QTLs conferring resistance to ascochyta blight in chickpea. Crop Sci 40:1606–1612

Schneider K, Weisshaar B, Borchardt DC, Salamani F (2001) SNP frequency and allele haplotype structure of *Beta vulgaris* expressed genes. Mol Breed 8:63–74

Sethy NK, Shokeen B, Edwards KJ, Bhatia S (2006) Development of microsatellite markers and analysis of intraspecific genetic variability in chickpea (*Cicer arietinum* L.) Theor Appl Genet 112:1416–1428

Sharma KD, Winter P, Muehlbauer FJ (2004) Molecular mapping of *Fusarium oxysporum* f. sp. *ciceris* race 3 resistance gene in chickpea. Theor Appl Genet 108:1243–1248

Tekeoglu M, Rajesh PN, Muehlbauer FJ (2002) Integration of sequence tagged microsatellite sites to the chickpea genetic map. Theor Appl Genet 105:847–854

Winter P, Pfaff T, Udupa SM, Hüttel B, Sharma PC, Sahi S, Arreguin-Espinoza R, Weigand F, Muehlbauer FJ, Kahl G (1999) Characterization and mapping of sequence-tagged microsatellite sites in the chickpea (*Cicer arietinum* L). Mol Gen Genet 262:90–101

Winter P, Benko-Iseppon HB, Ratnaparkhe M, Tullu A, Sonnante G, Ptaff T, Tekeoglu M, Santra D, Sant VJ, Rajesh PN, Kahl G, Muehlbauer FJ (2000) A linkage map of the chickpea (*Cicer arietinum* L.) genome based on recombinant inbred lines from a *C. arietinum* X *C. reticulatum* cross: localization of resistance genes for *fusarium* wilt races 4 and 5. Theor Appl Genet 101:1155–1163

Zhu YL, Song QJ, Hyten DL, Van Tassell CP, Matukumalli LK, Grimm DR, Hyatt SM, Fickus EW, Young ND, Cregan PB (2003) Single-nucleotide polymorphisms in soybean. Genetics 163:1123–1134